

# Approximating the largest eigenvalue of network adjacency matrices

Juan G. Restrepo,<sup>1,\*</sup> Edward Ott,<sup>2</sup> and Brian R. Hunt<sup>3</sup>

<sup>1</sup>*Center for Interdisciplinary Research in Complex Systems,  
Northeastern University, Boston, MA 02115, USA*

<sup>2</sup>*Department of Physics, Department of Electrical and Computer Engineering,  
and Institute for Research in Electronics and Applied Physics,  
University of Maryland, College Park, Maryland 20742, USA*

<sup>3</sup>*Department of Mathematics and Institute for Physical Science and Technology,  
University of Maryland, College Park, Maryland 20742, USA*

(Dated: February 1, 2008)

The largest eigenvalue of the adjacency matrix of a network plays an important role in several network processes (e.g., synchronization of oscillators, percolation on directed networks, linear stability of equilibria of network coupled systems, etc.). In this paper we develop approximations to the largest eigenvalue of adjacency matrices and discuss the relationships between these approximations. Numerical experiments on simulated networks are used to test our results.

PACS numbers: 05.45.-a, 05.45.Xt, 89.75.-k

## I. INTRODUCTION

In recent years, there has been much interest in the study of the structure of networks arising from real world systems [1]. Another concern has been dynamical processes taking place on networks, and the impact of network structure on such dynamics. The largest eigenvalue of the network adjacency matrix has emerged as a key quantity important for the study of a variety of different dynamical network processes. In particular, this is the case for synchronization of coupled heterogeneous oscillators [2], percolation on directed networks [3], linear stability of the fixed points of systems of network-coupled ordinary differential equations [4], and several other examples in physics and chemistry [5, 6]. In this paper we study methods of obtaining approximations to the maximum eigenvalue for the adjacency matrix for the case of large complex networks.

We consider a network as a directed graph with  $N$  nodes, and we associate to it an  $N \times N$  adjacency matrix whose elements  $A_{ij}$  are one if there is a directed link from  $i$  to  $j$  and zero otherwise ( $A_{ii} \equiv 0$ ). We denote the largest eigenvalue of  $A$  by  $\lambda$  (it may be shown that of all the eigenvalues of  $A$ , the one with largest magnitude is real and positive [6]). Furthermore, we note that it is often the case that the largest eigenvalue is well separated from the second largest eigenvalue.

The properties of  $\lambda$  have been studied in the context of small or regular graphs [5] and in classical Erdős-Rényi random graphs [7]. However, the structure of real world networks is usually more complex, as evidenced by the fact that the degree distribution in a large number of examples has been found to be highly heterogeneous (often following a power law [8]), where the *out-degree* and *in-degree* of a node  $i$  are defined by

$d_i^{out} = \sum_{j=1}^N A_{ij}$  and  $d_i^{in} = \sum_{j=1}^N A_{ji}$ . The ‘degree distributions’,  $\hat{P}(d^{in}, d^{out})$ ,  $P_{in}(d^{in}) = \int P(d^{in}, d^{out}) dd^{out}$ , and  $P_{out}(d^{out}) = \int P(d^{in}, d^{out}) dd^{in}$ , are defined as the probabilities that a randomly chosen node has degree  $d^{in}$  and/or  $d^{out}$ . If  $A = A^T$  we say the graph is undirected. For an undirected graph each link serves as both an in and out link for each of the two nodes it joins and for any node  $i$  we have  $d_i^{in} = d_i^{out} \equiv d_i$ . Thus in the undirected case we write  $P(d)$  to denote the corresponding degree distribution. The effect of the degree distribution on the largest eigenvalue of the adjacency matrix has been explored recently by Chung *et al.* [9], who considered a particular ensemble of random *uncorrelated*, undirected networks whose number of nodes  $N$  is large (see also [10, 11, 12]). Here, by uncorrelated we mean that we regard the network to be a random draw from some ensemble of networks for which the joint probability distribution of the node degrees  $Q(d_1, d_2, \dots, d_N)$  factors,  $Q(d_1, d_2, \dots, d_N) = P(d_1)P(d_2) \dots P(d_N)$ . Chung *et al.* found that in the limit  $N \rightarrow \infty$  these networks yield an expected largest eigenvalue that is determined by the ratio  $\hat{\lambda}$  of the second to first moment of the average degree distribution,

$$\hat{\lambda} = \langle d^2 \rangle / \langle d \rangle, \quad (1)$$

where  $\langle x \rangle = N^{-1} \sum_{i=1}^N x_i$ , and by the expected largest degree  $\bar{d}_{max}$ . Specifically, they found that

$$\lambda \approx \begin{cases} \hat{\lambda} & \hat{\lambda} > \bar{d}_{max} \log N, \\ \sqrt{\bar{d}_{max}} & \sqrt{\bar{d}_{max}} > \hat{\lambda} \log^2 N. \end{cases} \quad (2)$$

Some previous results for dynamical processes in networks have been stated in terms of the quantity  $\hat{\lambda}$ , for example, the synchronization threshold in the *mean field theory* of coupled oscillators in networks [2, 13, 14] and the network percolation and epidemic spreading thresholds [15, 16].

Real world networks often have some amount of degree-degree correlations [17], i.e., a node of a given degree is

---

\*Electronic address: juanga@neu.edu

more likely to be connected to nodes with particular degrees than would be expected on the basis of chance. Networks in which high degree nodes connect preferentially to high (low) degree nodes, and viceversa, are called *assortative* (*disassortative*). Such correlations can affect dynamical process on networks, as has been demonstrated for example in epidemic spreading models and percolation [18, 19, 20].

We also emphasize that the in- and out- degrees at a node can have different distributions [i.e.,  $P_{in}(d^{in}) \neq P_{out}(d^{out})$ ], as has been noted for some corporate information and genetic networks [21, 22], and that there are potential correlations between the in- and out- degrees at the *same* node which can also significantly affect the largest eigenvalue.

The rest of this paper is organized as follows. Section II reviews the characterization of degree correlations. Section III develops the theory of the maximum eigenvalue  $\lambda$  for the case of networks that satisfy a certain Markovian property. Some of the considerations of Sec. III are similar to theory in previous papers, where, however, those previous considerations were application specific, and the more general applicability to the largest eigenvalue was not apparent. Section IV considers results for numerical constructions of network adjacency matrices of different types.

## II. DEGREE-DEGREE CORRELATIONS

### A. In and out degree correlations

As we discussed in the introduction, we define  $\hat{P}(d^{in}, d^{out})$  as the probability that a randomly chosen node has in-degree  $d^{in}$  and out-degree  $d^{out}$ .

As we shall see later, for networks without neighbor degree-degree correlations, a first order approximation to the eigenvalue  $\lambda$ , generalizing Eq. (1), is given by

$$\hat{\lambda} = \langle d^{in} d^{out} \rangle / \langle d \rangle, \quad (3)$$

where we recall that  $\langle \cdot \rangle$  denotes an average over nodes, and  $\langle d \rangle$  means either  $\langle d_i^{in} \rangle$  or  $\langle d_i^{out} \rangle$  which are equal:  $\langle d_i^{in} \rangle = N^{-1} \sum_i d_i^{in} = N^{-1} \sum_{i,j} A_{ij} = \langle d_i^{out} \rangle$ .

If  $d_i^{in}$  and  $d_i^{out}$  are independent, the largest eigenvalue is approximately given by  $\hat{\lambda} = \langle d \rangle$ , and if  $d_i^{in}$  and  $d_i^{out}$  are perfectly correlated, so that  $d_i^{in} = d_i^{out} \equiv d_i$  (e.g., as in an undirected network), the largest eigenvalue is approximately  $\hat{\lambda} = \langle d^2 \rangle / \langle d \rangle$ . We see that correlations between  $d^{in}$  and  $d^{out}$  can crucially affect the eigenvalue, especially if the second moment of either degree distribution diverges with increasing  $N$ , while the first moment converges: in such a case independence leads to a finite eigenvalue estimate, and perfect correlation to a diverging eigenvalue estimate.

In order to quantify the correlations between  $d^{in}$  and  $d^{out}$  at a node, we define the node degree correlation

coefficient,

$$\eta \equiv \langle d^{in} d^{out} \rangle / \langle d \rangle^2. \quad (4)$$

Note that if there are no correlations,  $\eta = 1$ , and  $\eta$  is larger (smaller) than 1 for positive (negative) correlations. In terms of the correlation coefficient,  $\hat{\lambda}$  is given by

$$\hat{\lambda} = \eta \langle d \rangle. \quad (5)$$

For undirected networks  $A$  is a symmetric positive matrix, and thus its largest eigenvalue satisfies

$$\lambda \geq \frac{q^T A q}{q^T q}, \quad (6)$$

for any  $N$ -vector  $q$ . Choosing the components of  $q$  to be zero for nodes not connected to the node of largest degree  $d_{max}$ , to be one for nodes connected to the node of largest degree, and  $\sqrt{d_{max}}$  for the node of largest degree, Eq. (6) yields

$$\lambda \geq \sqrt{d_{max}}. \quad (7)$$

If  $\sqrt{d_{max}} > \hat{\lambda}$ , the mean-field approximation (3) must be incorrect. Thus, for undirected networks we use as a heuristic alternative to the mean-field approximation,

$$\lambda \approx \max\{\hat{\lambda}, \sqrt{d_{max}}\}, \quad (8)$$

which is consistent with both of the regimes considered in the rigorous result (2). (We remark, however, that Eq. (8) holds in principle only if the ratio  $\hat{\lambda}/\sqrt{d_{max}}$  or its inverse is large enough [see Eq. (2)]). One way of viewing Eq. (8) is that Eq. (6) implies that  $\lambda$  is at least as large as the maximum eigenvalue of any subnetwork of the original network (by a subnetwork we mean one obtained by deleting links of the original network). Considering a subnetwork consisting of the node of maximum degree and the nodes connecting to it, shown in Fig. 1(a), Eq. (7) corresponds to the fact that  $\sqrt{d_{max}}$  is the maximum eigenvalue of this star network. The regime  $\lambda \approx \sqrt{d_{max}}$  in Eq. (8) applies to networks whose largest eigenvalue is dominated by the node with largest degree.

We now contrast the above situation for undirected networks with what can happen for directed networks. We first note that the reasoning leading to (8) may not hold. For example, Eq. (6) no longer applies, and a node with many in and out links does not necessarily make a subnetwork with large eigenvalue. Regarding the latter point, consider Fig. 1(b) as compared to Fig. 1(a). This directed network has all its eigenvalues zero, because no pair of in and out links connect the same two nodes. Moreover, if one views  $d^{in}$  and  $d^{out}$  to diverge as  $N$  increases at a rate sufficiently more slowly than  $N$ , then, as  $N$  approaches infinity, the probability of an out link and an in link from the central node connecting to the same peripheral node goes to zero with increasing  $N$ , if these peripheral nodes are randomly chosen.

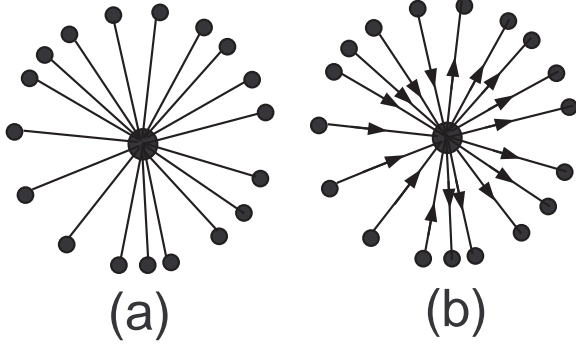


FIG. 1: Undirected (a) and directed (b) star networks illustrating the fact that, in the directed case, a node with many in and out links does not necessarily constitute a subnetwork with large eigenvalue (see text).

### B. Neighbor degree correlations

For our subsequent analysis it is useful to introduce an edge degree correlation coefficient  $\rho$  characterizing the correlation between the in-degree at node  $i$  and the out-degree at node  $j$ , where a directed link goes from  $i$  to  $j$ ,

$$\rho \equiv \langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e / \langle d_i^{\text{in}} \rangle_e \langle d_j^{\text{out}} \rangle_e, \quad (9)$$

where  $\langle Q_{ij} \rangle_e$  denotes an average over edges,

$$\langle Q_{ij} \rangle_e \equiv \sum_{i,j} A_{ij} Q_{ij} / \sum_{i,j} A_{ij}. \quad (10)$$

Using (10) we have

$$\begin{aligned} \langle d_i^{\text{in}} \rangle_e &= \sum_{i,j} A_{ij} d_i^{\text{in}} / \sum_{i,j} A_{ij} = \sum_i d_i^{\text{out}} d_i^{\text{in}} / \sum_{i,j} A_{ij} \\ &= \langle d^{\text{out}} d^{\text{in}} \rangle / \langle d \rangle, \end{aligned} \quad (11)$$

and so

$$\rho = \frac{\langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e}{\eta^2 \langle d \rangle^2}, \quad (12)$$

(Our definition of the edge degree correlation coefficient  $\rho$  is related to but slightly different from that of Newman [17].)

We note that the term  $\sum_{i,j} A_{ij} d_i^{\text{in}} d_j^{\text{out}}$  that appears in the definition of  $\langle d_i^{\text{in}} d_j^{\text{out}} \rangle_e$  is the number of directed paths of length three, i.e., the number of links of the form  $n \rightarrow i \rightarrow j \rightarrow m$ . Similarly,  $\langle d_i^{\text{in}} \rangle_e N \langle d \rangle = \langle d^{\text{out}} \rangle_e N \langle d \rangle = N \langle d^{\text{in}} d^{\text{out}} \rangle = \sum_k d_k^{\text{out}} d_k^{\text{in}} = \sum_{i,j,k} A_{ik} A_{kj}$  is the number of paths of length 2, and  $N \langle d \rangle$  is the number of paths

of length 1 (or the number of links). Accordingly, let us write  $\sum_{i,j} A_{ij} d_i^{\text{in}} d_j^{\text{out}} \equiv n_3$ ,  $N \langle d^{\text{in}} d^{\text{out}} \rangle \equiv n_2$ , and  $N \langle d \rangle \equiv n_1$ . With this notation, the coefficients  $\eta$  and  $\rho$  can be rewritten as

$$\eta = \frac{n_2}{n_1 \langle d \rangle}, \quad (13)$$

$$\rho = \frac{n_3 n_1}{n_2^2}, \quad (14)$$

As an example of this interpretation, we note that for networks with uncorrelated in- and out-degrees, the number of paths of length two is  $n_2 \approx n_1 \langle d \rangle$ , and so  $\eta \approx 1$ . For networks where there are no neighbor degree correlations, the number of paths of length three,  $n_3$ , can be obtained from the number of paths of length two,  $n_2$ , times the average branching ratio given by  $n_2/n_1$ , and thus  $\rho \approx 1$  for such networks.

## III. LARGEST EIGENVALUE OF MARKOVIAN NETWORKS

### A. Formulation

For generality, in this section we allow different types of network nodes, where we specify the node type by an index  $\sigma = 1, 2, \dots, M$ , where  $M$  is the total number of possible node types (e.g., in the case of social networks connecting people,  $\sigma$  might label sex, race, social class, etc.). Furthermore, we introduce the quantity

$$z = (d^{\text{in}}, d^{\text{out}}, \sigma), \quad (15)$$

which we refer to as the degree. With this definition, we use  $P(z)$  to denote the degree distribution, i.e.,  $P(z)$  is the probability that a randomly chosen node has degree  $z$ . This implies, for example, that

$$\sum_{d^{\text{in}}, d^{\text{out}}} P(z) = N_\sigma / N, \quad (16)$$

where  $N_\sigma$  is the number of nodes of type  $\sigma$  and  $N = \sum_\sigma N_\sigma$ . (While our numerical examples in Sec. IV are for the case of a single node type,  $M = 1$ , the subsequent considerations in the present section do not have this restriction.)

We consider a particular class of networks for which the only nontrivial correlations are between nodes that are directly connected by a single link. Such *Markovian* networks have been considered in previous works on epidemic spreading and percolation [15, 16, 18, 20]. Under this assumption, if we define  $P(z'|z)$  to be the probability that a node with degree  $z$  has an outgoing link to a node with degree  $z'$ , then if we choose a random outward path of length two from a node with degree  $z$ , the probability that the first hop ends on a node with degree  $z'$  and the second on a node with degree  $z''$  is, by this assumption,

$$P(z', z''|z) = P(z''|z')P(z'|z).$$

Let  $\psi_z^{(m)}$  be the expected number of directed paths of length  $m$  whose starting node has degree  $z$ . Using the assumption that the network is Markovian, we can express  $\psi_z^{(m+1)}$  in terms of  $\psi_z^{(m)}$  as

$$\psi_z^{(m+1)} = d^{out} \sum_{z'} P(z'|z) \psi_{z'}^{(m)}. \quad (17)$$

The number of paths of length  $m$  grows, in the limit of large  $m$ , as the largest eigenvalue  $\lambda$ ,  $\psi_z^{(m)} \sim \psi_z \lambda^m$  [5]. Therefore, we can associate to Eq. (17) the eigenvalue problem

$$\lambda_C \psi_z = d^{out} \sum_{z'} P(z'|z) \psi_{z'}, \quad (18)$$

and we seek its largest eigenvalue  $\lambda_C$  and its corresponding eigenfunction  $\psi_z$ . We consider  $\lambda_C$  to be an approximation to  $\lambda$  that is more accurate than the mean field result in that it includes correlations between connected nodes,  $\lambda \approx \lambda_C$ . Additionally, an approximation  $u_C$  to the right eigenvector  $u$  of  $A$  can be obtained in terms of  $\psi_z$  in a similar way:

$$(u_C)_i = \psi_{z_i}, \quad (19)$$

and an analogous expression holds for the left eigenvector  $v$  of  $A$ , using the left eigenvector of the matrix  $d^{out} P(z'|z)$  instead of  $\psi_z$ . We will refer to Eqs. (18) and (19) as the *Markovian approximation*.

### B. $\epsilon$ expansion

While in many cases Eq. (18) can be solved directly, it is also of interest to explore approximations to its solution. Thus, we will here expand (18) about a zero order approximation in which degree-degree correlations are neglected and in which the first order correction gives the perturbation to the uncorrelated case due to a small amount of degree-degree correlations. In particular, we will be interested in zeroth order and first order approximations to the largest eigenvalue of  $A$  and to the corresponding right and left eigenvectors  $u$  and  $v$ , where

$$Au = \lambda u \quad \text{and} \quad v^T A = \lambda v^T. \quad (20)$$

Following previously used terminology, we refer to the zeroth approximation as the *mean field theory*.

When there are no neighbor degree-degree correlations,  $P(z'|z)$  becomes independent of  $z$ :

$$P(z'|z) = \hat{P}(z') = (d^{in})' P(z') / \langle d \rangle, \quad (21)$$

where  $P(z)$  is the degree distribution. We will expand Eq. (18) about the uncorrelated case. For this purpose, we will write

$$P(z'|z) = \hat{P}(z') + \epsilon \delta P(z'|z), \quad (22)$$

where

$$\delta P(z'|z) = P(z'|z) - \hat{P}(z'), \quad (23)$$

and  $\epsilon$  is an expansion parameter that we formally consider small, although, in reality,  $\epsilon = 1$ . Introducing expansions for  $\lambda_C$  and  $\psi_z$ ,

$$\lambda_C = \hat{\lambda} + \epsilon \delta \lambda + \dots \quad (24)$$

$$\psi_z = \hat{\psi}_z + \epsilon \delta \psi_z + \dots \quad (25)$$

Expanding Eq. (18) to zero order in  $\epsilon$ , we obtain

$$\hat{\psi}_z = d^{out}, \quad (26)$$

and

$$\hat{\lambda} = \langle d^{in} d^{out} \rangle / \langle d \rangle. \quad (27)$$

Thus in the zeroth order approximation, the right eigenvector of  $A$  has components

$$u_i = d_i^{out}. \quad (28)$$

To obtain the left eigenvector, we follow the same steps as above but with  $A$  replaced by  $A^T$ . This interchanges the roles of  $d^{in}$  and  $d^{out}$ , thus yielding

$$v_i = d_i^{in} \quad (29)$$

for the left eigenvector.

Expanding to first order in  $\epsilon$ , we obtain

$$\hat{\lambda} \delta \psi_z + \delta \lambda \hat{\psi}_z = d^{out} \sum_{z'} \hat{P}(z') \delta \psi_{z'} + d^{out} \sum_{z'} \delta P(z'|z) \hat{\psi}_{z'}. \quad (30)$$

Multiplying by  $\hat{P}(z)$  and summing over  $z$  we obtain, after some simplification,

$$\delta \lambda = \left[ (\hat{\lambda} \langle d \rangle)^{-1} \sum_{z, z'} (d^{out})' d^{out} d^{in} P(z'|z) P(z) \right] - \hat{\lambda} \quad (31)$$

The probability that a randomly chosen edge starts at a node with degree  $z$  is  $\hat{P}(z) = \langle d \rangle^{-1} P(z) d^{out}$ , and, therefore, the term  $\langle d \rangle^{-1} \sum_{z, z'} (d^{out})' d^{out} d^{in} P(z'|z) P(z)$  is equal to

$$\sum_{z, z'} d^{in} (d^{out})' \tilde{P}(z') P(z'|z) = \langle d_i^{in} d_j^{out} \rangle_e. \quad (32)$$

To first order, therefore, we obtain

$$\lambda_C \approx \hat{\lambda} + \delta \lambda = \frac{\langle d_i^{in} d_j^{out} \rangle_e}{\hat{\lambda}} = \hat{\lambda} \rho, \quad (33)$$

where  $\rho$  is defined in Sec. II, and we call  $\lambda = \hat{\lambda} \rho$  the *linear approximation*.

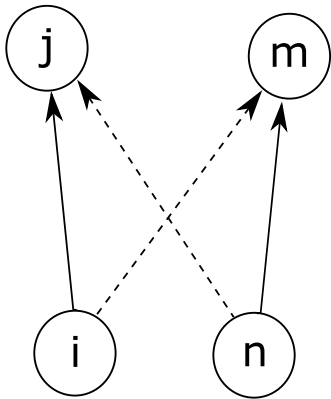


FIG. 2: Schematic representation of the rewiring algorithm. Edges  $i \rightarrow j$  and  $n \rightarrow m$  were chosen at random. When creating a network with assortative (disassortative) correlations, they are replaced with edges  $i \rightarrow m$  and  $n \rightarrow j$  if the resulting network has a larger (smaller) value of  $\rho$ .

Similarly,

$$\delta\psi_z = \hat{\lambda}^{-1} \sum_{z'} \delta P(z'|z)(d^{out})'. \quad (34)$$

We now briefly comment on the range of validity of the expansion results. As the rigorous result Eq. (2) shows, the mean field result for undirected networks is not valid when the network is dominated by the node with maximum degree. A network in which the maximum degree is too large compared with the bulk of the degree distribution can not satisfy Eq. (21). For example, the star network of Fig. 1(a) is very degree-degree correlated because all of its outer nodes connect only to the high degree hub node and, indeed, application of Eq. (21) to this simple star network yields an incorrect result. Our expansions about an uncorrelated network therefore implicitly assumed that this network was not dominated by the node with maximum degree. We speculate that directed networks might satisfy Eq. (27) with a less restrictive condition on the maximum degrees, as argued in Sec. II and Fig. 1. A more rigorous delineation of the range of validity of the Markovian approximation for correlated and/or directed networks along the lines of Ref. [9] is open for further research.

#### IV. NUMERICAL TESTS ON SIMULATED NETWORKS

In this section we numerically construct random networks and use them to compare with the predictions of Sec. III. We shall restrict ourselves to networks which are not dominated by nodes with large degree, as discussed at the end of Sec. IIIB. In order to study the variation of the largest eigenvalue  $\lambda$  of the adjacency matrix  $A$ , we first construct large ( $N \gg 1$ ) approxi-

mately uncorrelated networks with a given expected degree distribution using a generalization of the Random Graph model of Chung *et al.* [9]. First we associate to each node  $k$  target in and out degrees  $(\hat{d}_k^{in}, \hat{d}_k^{out})$ , where  $\sum_k \hat{d}_k^{in} = \sum_k \hat{d}_k^{out} = N\langle\hat{d}\rangle$ . Note that the choice of target node degrees can be made to correspond to a desired degree distribution (e.g., scale free, or Poisson). We then choose each element of the adjacency matrix  $A_{ij}$  randomly to be one with probability  $\hat{d}_i^{out} \hat{d}_j^{in} / (N\langle\hat{d}\rangle)$  and zero otherwise. This determines a network realization with in and out degrees  $d_i^{out} = \sum_j A_{ij}$  and  $d_j^{in} = \sum_i A_{ij}$ . In general,  $(d_k^{in}, d_k^{out})$  can be different from the target values  $(\hat{d}_k^{in}, \hat{d}_k^{out})$ . Nevertheless, with high probability, for large  $N$ , the resulting degree distribution  $P(d^{in}, d^{out})$  of this randomly chosen network will be approximately, in a suitable sense, the target distribution [9]. In particular, the moments  $\langle d^{in} d^{out} \rangle$  and  $\langle d \rangle$  will be approximately unchanged when calculated using either  $P$  or the target degrees. Furthermore,  $\rho \approx 1$  for  $N \gg 1$ . The network generated by this algorithm is directed (i.e.,  $A$  is asymmetric); symmetric networks can be generated by first considering only  $i < j$  and then setting  $A_{ji} = A_{ij}$ . Starting from an uncorrelated network generated by this algorithm, we then rewire the connections in such a way that the degree distribution is preserved. On doing so, the correlation coefficient changes and we calculate the largest eigenvalue for different values of  $\rho$ . The rewiring algorithm we use is a simplified version of that used in [19] and consists of the repeated iteration of the following steps:

1. Two edges are chosen at random. Assume one connects node  $n$  to node  $m$  and the other connects node  $i$  to node  $j$ .

2. Let

$$H(i, j; n, m) = d_n^{in} d_m^{out} + d_i^{in} d_j^{out} - d_n^{in} d_j^{out} - d_i^{in} d_m^{out}. \quad (35)$$

The two edges chosen in step 1 are replaced with two edges connecting node  $n$  to node  $j$  and node  $i$  to node  $m$  (see Fig. 2) if  $sH(i, j; n, m) < 0$ , and are left alone otherwise. Setting  $s = 1$  or  $-1$  we produce assortative or disassortative networks, respectively.

As we iterate from  $\rho = 1$  with  $s = 1$  ( $s = -1$ ),  $\rho$  steadily increases (decreases). Thus, we produce a sequence of networks with successively larger (smaller)  $\rho$ .

In Fig. 3 a) we show the largest eigenvalue  $\lambda$  (solid line), the linear approximation given by Eq. (33) (dashed line), and the Markovian approximation  $\lambda_C$  from Eq. (18) (boxes) as a function of the correlation coefficient  $\rho$  for an undirected network with power law degree distribution generated as described above with  $N = 25000$ ,  $\langle d \rangle = 100$ , and exponent  $\gamma = 2.5$ . In Fig. 3 b) we plot the same quantities for a directed network with  $N = 10000$ ,  $\langle d \rangle = 20$ , and exponent  $\gamma = 2.5$ . In these plots there is no discernible difference between the approximation  $\lambda_C$  and the actual values of  $\lambda$ . We observe that the largest

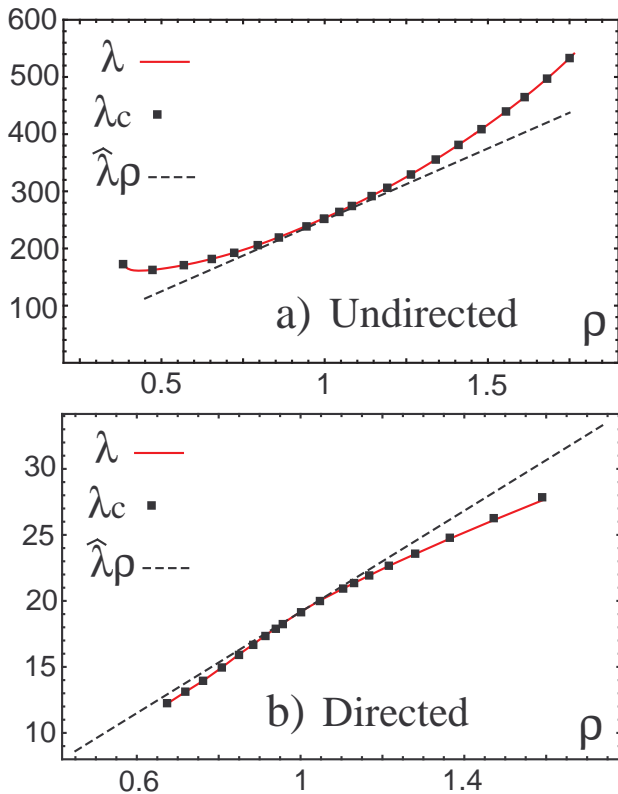


FIG. 3: Largest eigenvalue  $\lambda$  (solid line), linear approximation [Eq. (33), dashed line], and Markovian approximation  $\lambda_C$  [Eq. (18), boxes] for a) an undirected network with  $N = 25000$ ,  $\gamma = 2.5$  and  $\langle d \rangle = 100$ , and b) a directed network with  $N = 10000$ ,  $\gamma = 2.5$  and  $\langle d \rangle = 20$ .

eigenvalue depends strongly on the correlation coefficient: in the undirected case, it increases more than three times as  $\rho$  varies from 0.4 to 1.7. Also, we see that in these examples the linear approximation works for  $|\rho - 1| \lesssim 0.2$ , but fails for larger values of  $|\rho - 1|$ . In the undirected case,  $\lambda$  is larger than the linear approximation, which follows from Eq. (6) if we set  $q_i = d_i$ .

In Fig. 4 we show the eigenvector  $u_i$  for the network of Fig. 3 b) at  $\rho \approx 0.9$  plotted against the corresponding approximation  $(u_C)_i$ , Eq. (19) using an arbitrary scale. There is good agreement between the true value and its Markov estimate.

## V. CONCLUSION

In this paper we have considered several approximations to the largest eigenvalue of the adjacency matrix of large, directed networks. The mean field result (3) appears to apply well to networks whose neighboring nodes are uncorrelated in their degrees. The linear approximation (33) applies for sufficiently small correlation, while

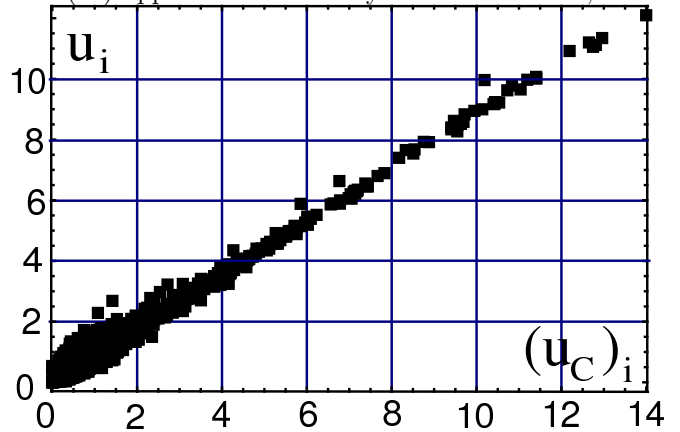


FIG. 4: Entries of the eigenvector  $u$  corresponding to  $\lambda$  versus the entries of its Markovian approximation  $u_C$  [Eq. (19)] in an arbitrary scale for the network of Fig. 3 b) at  $\rho \approx 0.9$ .

the Markov model (18) applies for arbitrarily strong degree correlations between neighbors. The price to be paid for a more refined approximation is the requirement of greater knowledge of the network (e.g., use of (18) requires knowledge of  $P(z|z')$  which is not required for the two other less refined approximations).

We caution that, although we have obtained good agreement between the theory and numerical results on simulated networks, this may not necessarily carry through for real networks encountered in practice. In particular, the Markov assumption of Eq. (17) may not always hold (e.g., due to community structure [23], clustering, or degree correlations extending over more than one link between nodes). This remains a topic for further study.

This work was supported by ONR (Physics), by the NSF (PHY 0456240 and DMS 0104-087), and by AFOSR FA95500410319.

- 
- [1] M. E. J. Newman, SIAM Review **45**, 167 (2003); A.-L. Barabási, and R. Albert, Rev. Mod. Phys. **74**, 47 (2002); S. Boccaletti *et al.*, Phys. Rep. **424**, 175 (2006).
  - [2] J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. E **71**, 036151 (2005); J. G. Restrepo, E. Ott, and B. R. Hunt, Chaos **16**, 015107 (2006); J. G. Restrepo, E. Ott, and B. R. Hunt, Phys. Rev. Lett. **96**, 254103 (2006).
  - [3] J. G. Restrepo, E. Ott, and B. R. Hunt, arXiv:0704.0491v1; B. Bollobas, C. Borgs, J. Chayes, and O. Riordan, arXiv:math/0701346v1.
  - [4] R. May, Nature **238**, 413 (1972); J. Feng, V. K. Jirsa, and M. Ding, Chaos **16** 015109 (2006); M. Brede and S. Sinha, eprint cond-mat/0507710.
  - [5] D. Cvetkovic and P. Rowlinson, Linear Multilinear Alge-

- bra **28**, 3 (1990).
- [6] C. R. MacCluer, SIAM Review **42**, 487 (2000).
  - [7] M. Krivelevich and B. Sudakov, Combinatorics, Probability and Computing **12**, 61 (2003).
  - [8] A.-L. Barabási, and R. Albert, Science **286**, 509 (1999).
  - [9] F. Chung, L. Lu and V. Vu, Proc. Natl. Acad. Sci., **100**, 6313 (2003).
  - [10] I. J. Farkas, I. Derényi, A.-L. Barabási, and T. Vicsek, Phys. Rev. E **64**, 026704 (2001).
  - [11] K.-I. Goh, B. Kahng, D. Kim, Phys. Rev. E **64**, 051903 (2001).
  - [12] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. E **68**, 046109 (2003).
  - [13] T. Ichinomiya, Phys. Rev. E **70**, 026116 (2004).
  - [14] D.-S. Lee, Phys. Rev. E **72**, 026208 (2005).
  - [15] R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Phys. Rev. Lett. **85**, 4626 (2000).
  - [16] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).
  - [17] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).
  - [18] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101(R) (2003).
  - [19] M. E. J. Newman, Phys. Rev. Lett. **89** 208701 (2002).
  - [20] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **66** 047104 (2002).
  - [21] D. Braha and Y. Bar-Yam, Phys. Rev. E **69**, 016113 (2004).
  - [22] N. Guelzim, S. Bottani, P. Bourguin, and François Képès, Nat. Genet. **31** 60 (2002).
  - [23] M. E. J. Newman and M. Girvan, Phys. Rev. E **69** 026113 (2004).